

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

FD 23/1497 71

us-09-446-543a-73.ra1

OM protein - protein search, using sw model

Run on: April 17, 2001, 15:39:49 ; Search time 61.54 Seconds  
 (without alignments)  
 6.868 Million cell updates/sec

Title: US-09-446-543A-73  
 Perfect score: 109  
 Sequence: 1 TPDINPAWYXXRGIREVGRFX 22

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 1857.57

Minimum DB seq length: 0  
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

**Database :**

Issued Patents\_AK: \*

1: /cgn2\_6/ptodata/2/1aa/5A..COMB.pep:\*

2: /cgn2\_6/ptodata/2/1aa/5B..COMB.pep:\*

3: /cgn2\_6/ptodata/2/1aa/6A..COMB.pep:\*

4: /cgn2\_6/ptodata/2/1aa/6B..COMB.pep:\*

5: /cgn2\_6/ptodata/2/1aa/pcus..COMB.pep:\*

6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	105	96.3	20	3 US-09-105-678A-46
2	105	96.3	21	3 US-09-105-678A-47
3	105	96.3	22	3 US-09-105-678A-48
4	105	96.3	31	3 US-09-105-678A-49
5	105	96.3	31	3 US-09-105-678A-43
6	105	96.3	32	3 US-09-105-678A-44
7	105	96.3	33	3 US-09-105-678A-45
8	104	95.4	20	3 US-09-105-678A-34
9	104	95.4	20	3 US-09-105-678A-40
10	104	95.4	21	3 US-09-105-678A-35
11	104	95.4	21	3 US-09-105-678A-41
12	104	95.4	22	3 US-09-105-678A-36
13	104	95.4	22	3 US-09-105-678A-42
14	104	95.4	31	3 US-09-105-678A-7
15	104	95.4	31	3 US-09-105-678A-8
16	104	95.4	31	3 US-09-105-678A-31
17	104	95.4	31	3 US-09-105-678A-28
18	104	95.4	31	4 US-09-172-353-4
19	95.4	32	3 US-09-105-678A-32	Sequence 32, Appl
20	95.4	32	3 US-09-105-678A-38	Sequence 32, Appl
21	104	95.4	33	3 US-09-105-678A-33
22	104	95.4	33	3 US-09-105-678A-39
23	103	94.5	21	3 US-09-105-678A-28
24	98	89.9	19	3 US-09-105-678A-30
25	93	85.3	29	3 US-09-105-678A-29
26	44	349	1	US-08-118-270-71
27	40.4	5	PCT-US93-08528-71	Sequence 71, Appl

**RESULT 1**

US-09-105-678A-46

Sequence 46, Application US/09105678A  
 Patent No. 6103822 on 8/15/2000

GENERAL INFORMATION:

APPLICANT: Sueyaga, Masato  
 APPLICANT: Moriya, Takeo  
 APPLICANT: Tanaka, Yoko

APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/105-678A

FILING DATE: 26-JUN-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: (7P-172118/1997)

FILING DATE: (27-JUN-1997)

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.  
 REGISTRATION NUMBER: 27,026  
 REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400  
 TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 20 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide

US-09-105-678A-46

Query Match 96.3%; Score 105; DB 3; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 1.5e-11;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

**RESULT 2**  
**US-09-105-678A-47**  
; Sequence 47, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105, 678A  
; FILING DATE: 26-JUN-1998  
; PRIORITY APPLICATION DATA:  
; REFERENCE/DOCKET NUMBER: 48466-342  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REFERENCE/DOCKET NUMBER: 27, 026  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-47  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-48  
; Query Match 96.3%; Score 105; DB 3; Length 21;  
; Best Local Similarity 90.0%; Pred. No. 1.5e-11; 18; Conservative  
; Matches 0; Mismatches 2; Indels 0; Gaps 0;  
; QY 1 TPDINPAWYXRGIRPVGRF 20  
; Db 1 TPDINPAWYASRGIRPVGRF 20  
; RESULT 4  
; US-09-105-678A-9  
; Sequence 9, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105, 678A  
; FILING DATE: 26-JUN-1998  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026  
 REFERENCE/DOCKET NUMBER: 48466-342  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-523-3400  
 TELEFAX: 617-523-6440  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 31 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-09-105-678A-9

RESULT 5  
 US-09-105-678A-43  
 ; Sequence 43, Application US/09105678A  
 ; Patent No. 6103882  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sueanaga, Masato  
 ; APPLICANT: Moriya, Takeo  
 ; APPLICANT: Tanaka, Yoko  
 ; APPLICANT: Nishimura, Osamu  
 ; ADDRESS: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 ; STREET: 130 Water Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/105,678A  
 FILING DATE: 26-JUN-1998  
 PRIORITY APPLICATION DATA:  
 PRIORITY NUMBER: JP 172118/1997  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/105,678A  
 FILING DATE: 26-JUN-1998  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: JP 172118/1997  
 FILING DATE: 27-JUN-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Conlin, David G.  
 REGISTRATION NUMBER: 27,026  
 REFERENCE/DOCKET NUMBER: 48466-342  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-523-3400  
 TELEFAX: 617-523-6440  
 INFORMATION FOR SEQ ID NO: 43:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 31 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-09-105-678A-43

RESULT 6  
 US-09-105-678A-44  
 ; Sequence 44, Application US/09105678A  
 ; Patent No. 6103882  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sueanaga, Masato  
 ; APPLICANT: Moriya, Takeo  
 ; APPLICANT: Tanaka, Yoko  
 ; APPLICANT: Nishimura, Osamu  
 ; ADDRESS: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 ; STREET: 130 Water Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/105,678A  
 FILING DATE: 27-JUN-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Conlin, David G.  
 REGISTRATION NUMBER: 27,026  
 REFERENCE/DOCKET NUMBER: 48466-342  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-523-3400  
 TELEFAX: 617-523-6440  
 INFORMATION FOR SEQ ID NO: 44:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 32 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-09-105-678A-44

RESULT 7  
 US-09-105-678A-45  
 ; Sequence 45, Application US/09105678A  
 ; Patent No. 6103882  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sueanaga, Masato  
 ; APPLICANT: Moriya, Takeo  
 ; APPLICANT: Tanaka, Yoko  
 ; APPLICANT: Nishimura, Osamu  
 ; ADDRESS: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 ; STREET: 130 Water Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/105,678A  
 FILING DATE: 27-JUN-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Conlin, David G.  
 REGISTRATION NUMBER: 27,026  
 REFERENCE/DOCKET NUMBER: 48466-342  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-523-3400  
 TELEFAX: 617-523-6440  
 INFORMATION FOR SEQ ID NO: 45:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 31 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-09-105-678A-45

Query Match 96.3%; Score 105; DB 3; Length 31;  
 Best Local Similarity 90.0%; Pred. No. 2.4e-11; 2;  
 Matches 18; Conservative 0; Mismatches 2;  
 Indels 0; Gaps 0;

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/105,678A  
 FILING DATE: 26-JUN-1998  
 PRIORITY APPLICATION DATA:

APPLICATION NUMBER: JP 172118/1997  
 FILING DATE: 27-JUN-1997  
 ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.  
 REGISTRATION NUMBER: 27,026  
 REFERENCE/DOCKET NUMBER: 48466-342  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-523-3400  
 TELEFAX: 617-522-6440

INFORMATION FOR SEQ ID NO: 34:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 20 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-09-105-678A-34

INFORMATION FOR SEQ ID NO: 45:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 33 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-09-105-678A-45

Query Match 95.3%; Score 105; DB 3; Length 33;  
 Best Local Similarity 90.0%; Pred. No. 2.6e-11;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAWYXRGTRPVGRF 20  
 QY 12 TPDINPAWYASRGIRGPVGRF 31

RESULT 8  
 US-09-105-678A-34

; Sequence 34, Application US/09105678A  
 ; Patent No. 6103882  
 ; GENERAL INFORMATION:

APPLICANT: Suenaga, Masato  
 APPLICANT: Moriya, Takeo  
 APPLICANT: Tanaka, Yoko  
 APPLICANT: Nishimura, Osamu  
 TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/105,678A  
 FILING DATE: 26-JUN-1998  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: JP 172118/1997  
 FILING DATE: 27-JUN-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Conlin, David G.  
 REGISTRATION NUMBER: 27,026  
 REFERENCE/DOCKET NUMBER: 48466-342  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-523-3400  
 TELEFAX: 617-522-6440

INFORMATION FOR SEQ ID NO: 40:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 20 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-09-105-678A-40

Query Match 95.4%; Score 104; DB 3; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 2.2e-11;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPPDINPAWYXRGIRGVGRF 20  
 US-09-105-678A-35  
 Sequence 35, Application US/09105678A  
 ; Patent No. 6103882  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sueaga, Masato  
 ; APPLICANT: Moriya, Takeo  
 ; APPLICANT: Tanaka, Yoko  
 ; APPLICANT: Nishimura, Osamu  
 ; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 ; NUMBER OF SEQUENCES: 52  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 ; STREET: 130 Water Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/105,678A  
 ; FILING DATE: 26-JUN-1998  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 172118/1997  
 ; FILING DATE: 27-JUN-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Conlin, David G.  
 ; REGISTRATION NUMBER: 27,026  
 ; REFERENCE/DOCKET NUMBER: 4B466-342  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-523-3400  
 ; TELEFAX: 617-523-4440  
 ; INFORMATION FOR SEQ ID NO: 41:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 21 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-09-105-678A-41

Query Match 95.4%; Score 104; DB 3; Length 21;  
 Best Local Similarity 90.0%; Pred. No. 2.3e-11; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPPDINPAWYXRGIRGVGRF 20  
 US-09-105-678A-36  
 Sequence 36, Application US/09105678A  
 ; Patent No. 6103882  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sueaga, Masato  
 ; APPLICANT: Moriya, Takeo  
 ; APPLICANT: Tanaka, Yoko  
 ; APPLICANT: Nishimura, Osamu  
 ; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 ; NUMBER OF SEQUENCES: 52  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 ; STREET: 130 Water Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/105,678A  
 ; FILING DATE: 26-JUN-1998  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 172118/1997  
 ; FILING DATE: 27-JUN-1997

RESULT 11  
 US-09-105-678A-41  
 ; Sequence 41, Application US/09105678A  
 ; Patent No. 6103882  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sueaga, Masato  
 ; APPLICANT: Moriya, Takeo  
 ; APPLICANT: Tanaka, Yoko  
 ; APPLICANT: Nishimura, Osamu  
 ; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 ; NUMBER OF SEQUENCES: 52  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 ; STREET: 130 Water Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/105,678A  
 ; FILING DATE: 26-JUN-1998  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 172118/1997  
 ; FILING DATE: 27-JUN-1997

## ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27, 026

REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 22 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLogy: linear

MOLECULE TYPE: peptide

US-09-105-678A-36

Query Match 95.4%; Score 104; DB 3; Length 22;  
 Best Local Similarity 90.0%; Pred. No. 2.4e-11; 0; Mismatches  
 Matches 18; Conservative 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYXRGTRPVGRF 20  
 Db 1 TPDINPAWYXRGTRPVGRF 20

Query Match 95.4%; Score 104; DB 3; Length 22;  
 Best Local Similarity 90.0%; Pred. No. 2.4e-11; 0; Mismatches  
 Matches 18; Conservative 0; Indels 0; Gaps 0;

RESULT 13  
 US-09-105-678A-42  
 ; Sequence 42, Application US/09105678A  
 ; Patent No. 6103882

GENERAL INFORMATION:  
 APPLICANT: Sueunaga, Masato  
 APPLICANT: Moriya, Takeo  
 APPLICANT: Tanaka, Yoko  
 APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 NUMBER OF SEQUENCES: 52  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/105,678A

FILING DATE: 26-TUN-1998  
 PRIORITY DATA:  
 APPLICATION NUMBER: JP 172118/1997

ATTORNEY/AGENT INFORMATION:  
 NAME: Conlin, David G.  
 REGISTRATION NUMBER: 27, 026

REFERENCE/DOCKET NUMBER: 48466-342  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-523-3400  
 TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 31 amino acids  
 TYPE: amino-acid

STRANDEDNESS:

TOPOLogy: linear

MOLECULE TYPE: peptide

US-09-105-678A-7

RESULT 15  
 US-09-105-678A-8  
 ; Sequence 8, Application US/09105678A  
 ; Patent No. 6103882

GENERAL INFORMATION:  
 APPLICANT: Sueunaga, Masato  
 APPLICANT: Moriya, Takeo  
 APPLICANT: Tanaka, Yoko  
 APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 NUMBER OF SEQUENCES: 52  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

Best Local Similarity 90.0%; Pred. No. 2.4e-11;  
 Matches 18; Conservative 0; Mismatches 0;  
 Indels 0; Gaps 0;

Qy 1 TPDINPAWYXRGTRPVGRF 20  
 Db 1 TPDINPAWYXRGTRPVGRF 20

RESULT 14  
 US-09-105-678A-7  
 ; Sequence 7, Application US/09105678A  
 ; Patent No. 6103882

GENERAL INFORMATION:  
 APPLICANT: Sueunaga, Masato  
 APPLICANT: Moriya, Takeo  
 APPLICANT: Tanaka, Yoko  
 APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 NUMBER OF SEQUENCES: 52  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/105,678A

FILING DATE: 26-TUN-1998  
 PRIORITY DATA:  
 APPLICATION NUMBER: JP 172118/1997

ATTORNEY/AGENT INFORMATION:  
 NAME: Conlin, David G.  
 REGISTRATION NUMBER: 27, 026

REFERENCE/DOCKET NUMBER: 48466-342  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-523-3400  
 TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 31 amino acids  
 TYPE: amino-acid

STRANDEDNESS:

TOPOLogy: linear

MOLECULE TYPE: peptide

US-09-105-678A-7

RESULT 15  
 US-09-105-678A-8  
 ; Sequence 8, Application US/09105678A  
 ; Patent No. 6103882

GENERAL INFORMATION:  
 APPLICANT: Sueunaga, Masato  
 APPLICANT: Moriya, Takeo  
 APPLICANT: Tanaka, Yoko  
 APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 NUMBER OF SEQUENCES: 52  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/105,678A

FILING DATE: 26-TUN-1998  
 PRIORITY DATA:  
 APPLICATION NUMBER: JP 172118/1997

ATTORNEY/AGENT INFORMATION:  
 NAME: Conlin, David G.  
 REGISTRATION NUMBER: 27, 026

REFERENCE/DOCKET NUMBER: 48466-342  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-523-3400  
 TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 42:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 22 amino acids  
 TYPE: amino acid

STRANDEDNESS:

TOPOLogy: linear

MOLECULE TYPE: peptide

Query Match 95.4%; Score 104; DB 3; Length 22;

NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE:

FLOPPY DISK

COMPUTER:

IBM PC COMPATIBLE

OPERATING SYSTEM:

PC-DOS/MS-DOS

SOFTWARE:

PATENT RELEASE #1.0, VERSION #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER:

US/09/1105,678A

FILING DATE:

26-JUN-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

JP 17211871997

FILING DATE:

27-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME:

CONLIN, DAVID G.

REGISTRATION NUMBER:

27,026

REFERENCE/DOCKET NUMBER:

48466-342

TELECOMMUNICATION INFORMATION:

TELEPHONE:

617-523-3400

TELEFAX:

617-523-6440

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH:

31 amino acids

TYPE:

AMINO ACID

STRANDEDNESS:

TOPOLOGY:

LINEAR

MOLECULE TYPE:

peptide

US-09-105-678A-8

Query Match 95.4%; Score 104; DB 3; Length 31;  
Best Local Similarity 90.0%; Fred. No. 3.5e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAWYXRGTRPVGRF 20  
Db 12 TPDINPAWYTGRCRGPVGRF 31

Search completed: April 17, 2001, 15:39:49  
Job time: 318 sec

**THIS PAGE BLANK (USPTO)**

us-09-446-543a-73.rag

1-15 Application

137

Page 1

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on:

April 17, 2001, 15:38:42 ; Search time 116.94 Seconds

(without alignments)

10.754 Million cell updates/sec

Title: US-09-446-543a-73  
Perfect score: 109  
Sequence: TPDINPAWYXRGIRPVGRFXX 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters:

390729

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_0401:\*

1: /SIDS1/gcdata/geneseq/geneseq/geneseq/geneseqp/AA1980.DAT:\*

2: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1981.DAT:\*

3: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1982.DAT:\*

4: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1983.DAT:\*

5: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1984.DAT:\*

6: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1985.DAT:\*

7: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1986.DAT:\*

8: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1987.DAT:\*

10: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1988.DAT:\*

11: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1989.DAT:\*

12: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1990.DAT:\*

13: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1991.DAT:\*

14: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1992.DAT:\*

15: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1993.DAT:\*

16: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1994.DAT:\*

17: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1995.DAT:\*

18: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1996.DAT:\*

19: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1998.DAT:\*

20: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA1999.DAT:\*

21: /SIDS1/gcdata/geneseq/geneseq/geneseqp/AA2000.DAT:\*

22: /SIDS1/gcdata/geneseq/geneseqp/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

## ALIGNMENTS

RESULT	1	
W31394		W31394 standard; Peptide; 20 AA.
XX		
W31394		AC
XX		
W31394		DT 06-APR-1998 (first entry)
XX		DE Human type G protein-coupled receptor ligand fragment 4.
XX		KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator; pituitary; central nervous system; pancreas; prophylactic; therapeutic agent.
XX		KW
OS		Homo sapiens.
XX		
PN		W0972436-A2.
XX		
PD		10-JUL-1997.
XX		
PF		26-DEC-1996; 96WO-JP03921.
XX		
PR		18-SEP-1996; 96JP-0246573.
PR		28-DEC-1995; 95JP-0343371.
PR		15-MAR-1995; 96JP-0059119.
PR		12-AUG-1996; 96JP-0211805.
XX		
PA		(TAKE ) TAKEDA CHEM IND LTD.
XX		
FUJII R, FUKUSUMI S, HABATA Y, HINUMA S, HOSOYA M;		
PI		
KAWAMATA Y, KITADA C;		
XX		
DR		WPI; 1997-363672/33.
DR		N-PSDB; V02431.

Result	No.	Score	Query Match Length	DB ID	Description
1	< 105	96.3	20	18 W31394	Human type G prote
2	< 105	96.3	20	20 W97235	Human type ligand
3	105	96.3	20	21 B10365	Human oxytocin sec
4	105	96.3	20	21 Y49294	19P2 ligand peptid
5	105	96.3	21	18 W31395	Human type G prote
6	105	96.3	21	21 B10366	Human oxytocin sec
7	105	96.3	22	18 W31396	Human type G prote
8	105	96.3	22	21 B10367	Human oxytocin sec
9	105	96.3	31	18 W31391	Human type G prote
10	105	96.3	31	20 W87235	Human type ligand
11	105	96.3	31	20 W87615	Human 19P2 ligand.

XX  
 XX Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 PT  
 XX  
 PS Claim 2; Page 185; 258pp; English.  
 XX  
 CC This sequence represents a peptide fragment from a novel human type  
 ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
 sequence represented in W31390 and is used in an assay to monitor ligand  
 binding to the G protein-coupled receptor protein. Pharmaceutical  
 compositions containing this ligand may be used as a pituitary function  
 modulator, a central nervous system modulator or a pancreatic function  
 modulator. This ligand could have specific applications as a prophylactic  
 or therapeutic agent for dementia, depression, hypokinetic syndrome,  
 disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
 growth hormone secretory disease, hyper- and polyphagia, hyperlipidemia,  
 hypercholesterolemia, hyperglyceraidaemia, hyperprolactinaemia, diabetes,  
 cancer, pancreatitis, renal disease, Turner's syndrome, asthma,  
 rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,  
 amyotrophic lateral sclerosis, acute myocardial infarction, infertility,  
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,  
 osteoporosis and/or oligogalactia. Assays can also be developed to screen  
 CC compounds which are capable of altering the binding activity of the  
 CC ligand affecting activation of the G protein-coupled receptor protein.  
 XX  
 SQ Sequence 20 AA:

Query Match 96.3%; Score 105; DB 18; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 1.3e-11;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy . 1 TPDINPAWYXRGIRPVGRF 20  
 Db 1 tpdinpawysrgirpvgrf 20

RESULT 2  
 W97236

ID W97236 standard; peptide; 20 AA.

XX  
 AC W97236;

XX  
 DT 06-MAY-1999 (first entry)

XX  
 DE Human type ligand polypeptide fragment.

XX  
 KW Rat type ligand; modulation; prolactin secretion;

XX  
 KW G protein-coupled receptor; GPCR; hypoovarianism; gonecyst carcinogenesis;

XX  
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;

XX  
 KW pituitary adenomatosis; brain tumour; emenopath; autoimmune disease;

XX  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;

XX  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;

XX  
 KW contraceptive; placental function; choriocarcinoma; hydatid mole;

XX  
 KW irrigation; mole; abortion; unthrifly fetus; abnormal saccharometabolism;

XX  
 OS Homo sapiens.

PN W09858962-A1.

XX  
 PR 30-DEC-1998.

XX  
 PP 22-JUN-1998; 98WO-JP02765.

XX  
 PR 23-JUN-1997; 97JP-0165437.

XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.

XX  
 PL Fujii, R, Hinuma, S, Kawamata, Y, Matsumoto, H;

XX  
 DR WPI; 1999-105614/09.

XX  
 XX Ligand peptide for G protein-coupled receptor - acts by modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX  
 PS Claim 3; Page 166; 241pp; English.

XX  
 XX Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX  
 PS Claim 3; Page 166; 241pp; English.

The present sequence represents a human type ligand fragment. It  
 is used in the course of the invention. The specification describes  
 an agent for modulating Prolactin secretion which comprises a  
 ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 protein. The agents for promoting prolactin secretion can be used for  
 treating or preventing hypopituitarism, gonostic cacogenesis, menopausal  
 syndrome, euthyroid or hypometabolism. They can be used for promoting  
 lactation in a domestic mammal and as an aphrodisiac. The agents for  
 inhibiting prolactin secretion can be used for treating or preventing  
 pituitary adenomatosis, brain tumour, emenopathy, autoimmune disease,  
 prolatinoma, infertility, impotence, amenorrhea, galactorrhea,  
 acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
 Forbes-Albright syndrome, Lymphoma, Sheehan syndrome or dyzoospermia.  
 The inhibitory agents can also be used as contraceptives. The agents for  
 modulating placental function can be used for treating or preventing  
 choriocarcinoma, hydatid mole, interruption mole, abortion, unthrifly fetus,  
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX  
 Sequence 20 AA:

Query Match 96.3%; Score 105; DB 20; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 1.3e-11;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy . 1 TPDINPAWYXRGIRPVGRF 20  
 Db 1 tpdinpawysrgirpvgrf 20

RESULT 3  
 B10365

ID B10365 standard; peptide; 20 AA.

XX  
 AC B10365;

XX  
 DT 24-NOV-2000 (first entry)

XX  
 DE Human oxytocin secretion promoting Peptide SEQ ID NO: 35.

XX  
 KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;

XX  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;

XX  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;

XX  
 KW veterinary medicine; milk production.

XX  
 OS Homo sapiens.

PN W020038704-A1.

XX  
 PD 06-JUL-2000.

XX  
 PR 22-DEC-1999; 99WO-JP07199.

XX  
 PR 25-DEC-1998; 98JP-0369585.

XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.

XX  
 PI Matsumoto, H, Kitada, C, Hinuma, S;

XX  
 DR WPI; 2000-452298/39.

XX  
 XX Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine

PS Disclosure; Page 63; 72pp; Japanese.  
 XX  
 CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.,  
 CC weak Pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter.  
 XX  
 SQ Sequence 20 AA;  
 Query Match 96.3%; Score 105; DB 21; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 1.3e-11; Indels 0; Gaps 0;  
 Matches 18; Conservative 0; Mismatches 2;  
 QY 1 TPDINPAWYXRGIRPVGRF 20  
 Db 1 tpdinpawysrgirpvgrf 20  
 RESULT 4  
 Y49294 standard; peptide; 20 AA.  
 ID Y49294  
 XX  
 AC Y49294;  
 XX  
 DT 22-FEB-2000 (first entry)  
 DE 19P2 ligand peptide fragment.  
 KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
 KW OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Modified-site 20 /note= "C-terminal amide"  
 XX  
 PN WO960112-A1.  
 XX  
 PD 25-NOV-1999.  
 XX  
 PT 20-MAY-1999; 99WO-JP02650.  
 XX  
 PR 21-MAY-1998; 98JP-0140293.  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PT Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
 PI Kawamoto Y, Kitada C;  
 XX  
 DR WPI: 1997-363672/33.  
 XX  
 N-PSDB; V02432.  
 PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Claim 2; Page 186; 258pp; English.  
 XX  
 CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 34 to 54 of the  
 CC sequence represented in W31390 and is used in an assay to monitor ligand  
 CC binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a prophylactic  
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,  
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,  
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,  
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,  
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,  
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,  
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,  
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen  
 CC compounds which are capable of altering the binding activity of the  
 CC ligand affecting activation of the G protein-coupled receptor protein.  
 XX  
 SQ Sequence 21 AA;  
 Query Match 96.3%; Score 105; DB 21; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 1.3e-11; Indels 0; Gaps 0;  
 Matches 18; Conservative 0; Mismatches 2;  
 QY 1 TPDINPAWYXRGIRPVGRF 20  
 Db 1 tpdinpawysrgirpvgrf 20  
 RESULT 5  
 W31395 standard; Peptide; 21 AA.  
 ID W31395  
 XX  
 AC W31395;  
 XX  
 DT 06-APR-1998 (first entry)  
 DE Human type G protein-coupled receptor ligand fragment 5.  
 KW G protein-coupled receptor; ligand binding; pharmaceutical;  
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
 KW therapeutic agent.  
 XX  
 PT Homo sapiens.  
 XX  
 PN W09724436-A2.  
 XX  
 PD 10-JUL-1997.  
 XX  
 PT 26-DEC-1996; 96WO-JP03821.  
 XX  
 PR 18-SEP-1996; 96JP-0246573.  
 PR 28-DEC-1995; 95JP-0343171.  
 PR 15-MAR-1996; 96JP-0059419.  
 PR 12-AUG-1996; 96JP-0211805.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PT Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
 PI Kawamoto Y, Kitada C;  
 XX  
 DR WPI: 1997-363672/33.  
 XX  
 N-PSDB; V02432.  
 PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Claim 2; Page 186; 258pp; English.  
 XX  
 CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 34 to 54 of the  
 CC sequence represented in W31390 and is used in an assay to monitor ligand  
 CC binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a prophylactic  
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,  
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,  
 CC hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes,  
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,  
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,  
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,  
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,  
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen  
 CC compounds which are capable of altering the binding activity of the  
 CC ligand affecting activation of the G protein-coupled receptor protein.  
 XX  
 SQ Sequence 21 AA;

Query Match 96.3%; Score 105; DB 18; Length 21;  
 Best Local Similarity 90.0%; Pred. No. 1.4e-11; Mismatches 2;  
 Matches 18; Conservative 0; Gaps 0; Indels 0;

Qy 1 TPDINPAWYXRGIRPVGRF 20  
 Db 1 tpdinpawyasrgirpvgrf 20

RESULT 6

ID B10366 standard; peptide; 21 AA.  
 XX  
 AC B10366;  
 XX  
 DT 24-NOV-2000 (first entry)  
 XX  
 DE Human oxytocin secretion promoting peptide SEQ ID NO: 36.  
 KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.  
 XX  
 OS Homo sapiens.  
 XX  
 WO200038704-A1.  
 XX  
 PR 06-JUL-2000.  
 XX  
 PE 22-DEC-1999; 99WO-JP07199.  
 XX  
 PN WO200038704-A1.  
 XX  
 PD 12-AUG-1996; 96JP-0211805.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PT Matsumoto H, Kitada C, Hinuma S;  
 XX  
 DR WPI; 2000-452298/39.  
 XX  
 PT Physiologically-active polypeptide recognized as ligand by G protein-coupled receptor protein, for promoting secretion of oxytocin for diseases relating to oxytocin secretion and in veterinary medicine  
 XX  
 PS Disclosure; Page 63; 72pp; Japanese.  
 XX  
 This invention describes a novel oxytocin secretion-regulating agent which contains a ligand peptide or its salt for the G protein-coupled receptor protein. It is useful in the form of drugs for ameliorating, preventing and treating diseases relating to oxytocin secretion e.g. weak pains and atonic bleeding, before and after expulsion of placenta, uterine recovery failure, caesarean section, stoppage of artificial fertilization or galactostasis and is also applicable in veterinary medicine for promoting milk production in cow, goat and pig. This sequence represents a human peptide which acts as an oxytocin secretion promoter.  
 XX  
 Sequence 21 AA;  
 SQ Query Match 96.3%; Score 105; DB 18; Length 22;  
 Best Local Similarity 90.0%; Pred. No. 1.5e-11; Mismatches 2;  
 Matches 18; Conservative 0; Gaps 0; Indels 0;

Qy 1 TPDINPAWYXRGIRPVGRF 20  
 Db 1 tpdinpawyasrgirpvgrf 20

RESULT 8

ID B10367 standard; peptide; 22 AA.  
 XX  
 W31396

AC B10367; PN WO9724436-A2.  
 XX DT 24-NOV-2000 (first entry) XX 10-JUL-1997.  
 XX DE Human oxytocin secretion promoting peptide SEQ ID NO: 37. XX  
 XX KW Human; oxytocin secretion promoter; G protein-coupled receptor protein; treatment; disease; pain; atonic bleeding; uterine recovery failure; cow; caesarean section; artificial fertilization; galactostasis; goat; pig; veterinary medicine; milk production.  
 XX OS Homo sapiens.  
 XX PN WO200038704-A1.  
 XX PD 06-JUL-2000.  
 XX PR 25-DEC-1998; 98JP-0369585.  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX PI Matsumoto H, Kitada C, Hinuma S:  
 XX DR WPI; 2000-452298/39.  
 XX PT Physiologically-active polypeptide recognized as ligand by G protein-coupled receptor protein, for promoting secretion of oxytocin, as drugs for diseases relating to oxytocin secretion and in veterinary medicine -  
 PT Disclosure; Page 64; 72pp; Japanese.  
 XX CC This invention describes a novel oxytocin secretion-regulating agent which contains a ligand peptide or its salt for the G protein-coupled receptor protein. It is useful in the form of drugs for ameliorating, preventing and treating diseases relating to oxytocin secretion e.g. weak pains and atomic bleeding, before and after expulsion of placenta, uterine recovery failure, caesarean section, stoppage of artificial fertilization or galactostasis and is also applicable in veterinary medicine for promoting milk production in cow, goat and pig. This sequence represents a human peptide which acts as an oxytocin secretion promoter.  
 CC SQ sequence 22 AA;  
 Query Match 96.3%; Score 105; DB 21; Length 22;  
 Best Local Similarity 90.0%; Pred. No. 1.5e-11; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 TPDINPAWXXRGRRPVGRF 20  
 Db 1 tpdinpawyasrgirpvgrf 20  
 RESULT 9  
 ID W31391 standard; Peptide: 31 AA.  
 AC W31391;  
 XX 06-APR-1998 (first entry)  
 XX DE Human type G protein-coupled receptor ligand fragment 1.  
 XX KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator; pituitary; central nervous system; pancreas; prophylactic; therapeutic agent.  
 XX OS Homo sapiens.  
 XX  
 RESULT 10  
 ID W97235 standard; peptide; 31 AA.  
 AC W97235;  
 XX DT 06-MAY-1999 (first entry)  
 XX DE Human type ligand polypeptide fragment.  
 XX KW Rat type ligand; modulation; prolactin secretion; G protein-coupled receptor; GPCR; hypoovariatism; gonocyst carcinogenesis; menopausal syndrome; euthyroid; hypometabolism; lactation; pituitary adenomatosis; brain tumour; emmenoptysis; autoimmune disease; prolactinoma; infertility; impotence; amenorrhea; galactorrhea; acromegaly; Chiari Frommel syndrome; Argonz del Castillo syndrome; Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia; contraceptive; placental function; choriocarcinoma; hydatid mole;





XX  
 CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 23 to 54 of the  
 CC sequence represented in W11390 and is used in an assay to monitor ligand  
 CC binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a prophylactic  
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,  
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,  
 CC hypercholesterolaemia, hyperglycaidaemia, hyperprolactinaemia, diabetes,  
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,  
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,  
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,  
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,  
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen  
 CC compounds which are capable of altering the binding activity of the  
 CC ligand affecting activation of the G protein-coupled receptor protein.  
 XX

SQ Sequence 32 AA;

Query Match

Best Local Similarity 96.3%; Score 105; DB 18; Length 32;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TPDINPAWXXRGIRPVGRF 20

Db 12 tpdinpawyssrgirgvgrf 31

Search completed: April 17, 2001, 15:38:42  
 Job time: 349 sec

CC preventing and treating diseases relating to oxytocin secretion e.g.  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter.

XX Sequence 32 AA;

Query Match

95.3%; Score 105; DB 21; Length 32;

Best Local Similarity 90.0%; Pred. No. 2.2e-11; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TPDINPAWXXRGIRPVGRF 20

Db 12 tpdinpawyssrgirgvgrf 31

RESULT 15

B10353 standard: peptide; 32 AA.

ID B10363;

XX DT 24-NOV-2000 (first entry)

DE Human oxytocin secretion promoting peptide SEQ ID NO: 33.

KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.

OS Homo sapiens.

XX PN WO200038704-A1.

XX PD 06-JUL-2000.

XX PP 22-DEC-1999; 99WO-JP07199.

XX PR 25-DEC-1998; 98JP-0369585.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Matsumoto H, Kitada C, Hinuma S;

XX DR WPI; 2000-452298/39.

PT Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine

XX Disclosure: Page 62; 72pp; Japanese.

XX CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 receptor protein. It is useful in the form of drugs for ameliorating,

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on:

April 17, 2001, 15:45:59 ; Search time 70.08 Seconds

{without alignments)  
21.574 Million cell updates/sec

Title: US-09-446-543A-73  
Perfect score: 109  
Sequence: 1 TPDINPAWYXXRGIRGVGRFXX 22

scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR67:  
1: pir1:  
2: pir2:  
3: pir3:  
4: pir4:  
\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

8

Result No.	Score	Query Match Length	DB ID	Description
1	49.5	45.4	F83376	conserved hypothetical protein PA2151 [imported] - <i>Pseudomonas aeruginosa</i> (strain PAO
2	49	45.0	T47959	C;Species: <i>Pseudomonas aeruginosa</i> C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
3	46	42.2	H82852	R;Stover, C.K.; Pham, X.Q.; Eswin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.I.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
4	45	41.3	T21969	; Lory, S.; Olson, M.V. Nature 405, 959-964, 2000
5	44	40.4	250	G833400
6	44	40.4	284	F71015
7	43.5	40.9	1501	T45623
8	43	39.4	220	C83292
9	43	39.4	232	G75608
10	43	39.4	309	T32376
11	43	39.4	376	2 A48197
12	43	39.4	376	B48197
13	43	39.4	476	2 G64720
14	43	39.4	719	2 S61046
15	43	39.4	1295	T16859
16	42.5	39.0	443	2 T21499
17	42	38.5	428	2 F81694
18	41.5	38.1	345	2 D84012
19	41	37.6	128	2 S76955
20	41	37.6	284	2 A75117
21	41	37.6	338	2 T20100
22	41	37.6	342	2 B64395
23	41	37.6	343	2 T46534
24	41	37.6	347	2 H64371
25	41	37.6	986	1 OIURGA
26	41	37.6	2155	2 T30197
27	40.5	37.2	779	2 T49717
28	40.5	37.2	1540	2 T45619
29	40	36.7	184	2 T35841

## ALIGNMENTS

30	40	36.7	324	2 T35901	probable arac fam1
31	40	36.7	329	2 H70744	hypothetical prote
32	40	36.7	341	2 T35426	probable oxidoredu
33	40	36.7	359	2 T40084	pww domain prote
34	40	36.7	390	2 G82844	cysteine synthase
35	40	36.7	430	1 B69009	conserved hypothet
36	40	36.7	462	2 T00708	violaxanthin de-ep
37	40	36.7	468	2 C83160	nitrile extrusion
38	40	36.7	546	2 A32260	cholesterol oxidase
39	40	36.7	581	2 A42743	pol polyprotein -
40	40	36.7	843	1 G8WKK	pol polyprotein -
41	40	36.7	1196	1 GINAVV	pol polyprotein -
42	40	36.7	1196	1 GINAVV	probable polypeptid
43	40	36.7	1762	2 T03222	probable polypeptid
44	40	36.7	2100	2 T03223	elastic titin - hu
45	40	36.7	7962	2 I38346	

Query Match 45.0%; Score 49; DB 2; Length 790;  
 Best Local Similarity 47.4%; Pred. No. 37; Mismatches 9; Conservatve 3; Indels 7; Gaps 0; Matches 9;

Qy 2 PDINPAWYXKGIRPGRF 20  
 Db 366 PPNNPRTYGSRGLOPHGRW 384

RESULT 3

H82852 hydroxybenzoate octaprenyltransferase xF0068 [imported] - *Xylella fastidiosa* (strain 9as

C; Species: *Xylella fastidiosa*  
 C; Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 17-Nov-2000  
 C; Accession: H82852  
 R; anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen  
 Nature 406, 151-157, 2000  
 A; Title: The genome sequence of the Plant Pathogen *Xylella fastidiosa*.  
 A; Reference number: A82535; MUID:20365717  
 A; Note: for a complete list of authors see reference number A59328 below  
 A; Status: Preliminary  
 A; Molecule type: DNA  
 A; Residues: 1-333 <SM>  
 A; Cross-references: GB:AE003860; GB:AE003849; NID:9104830; PIDN:AAF82881.1; GSDB:GN001  
 A; Experimental source: strain 9as  
 R; Simpson, A.J.G.; Reinach, F.C.; Arredondo, P.; Abreu, F.A.; Acencio, M.; Alwarenga, R.; A  
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrasco, D.M.; Carrer, H  
 as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
 submitted to Genbank, June 2000  
 A; Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.H.; Kuramae, E.E.; Lairig  
 Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, H  
 A; Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracola, E.C.; Miyaki, C.Y.;  
 Rodrigues, V.; Rosa, A.J. de M.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.P.;  
 Rodrigues, V.; Rosa, A.J. de M.; de Sa, R.G.; Santelli, R.V.; Savasak  
 A; Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
 M.; Tsuchiya, M.H.; Vallada, H.; van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
 A; Reference number: A59328  
 A; Contents: annotation  
 A; Genetics:  
 A; Gene: xF0068  
 C; Superfamily: 4-hydroxybenzoate octaprenyltransferase

Query Match 42.2%; Score 46; DB 2; Length 333;  
 Best Local Similarity 53.3%; Pred. No. 4.9; Mismatches 8; Conservatve 2; Indels 5; Gaps 0; Matches 8;

Qy 4 INPAWYXKGIRPGV 18  
 Db 54 LDPYWKLARGDRPG 68

RESULT 4

T21969 hypothetical protein F38E11.7 - *Caenorhabditis elegans*  
 C; Species: *Caenorhabditis elegans*  
 C; Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C; Accession: T21969  
 R; Matthews, P.  
 A; Status: preliminary; translated from GB/EMBL/DDJB  
 A; Molecule type: DNA  
 A; Residues: 1-767 <WIL>  
 A; Cross-references: EMBL:268342; PIDN:CAA92775.1; GSPDB:GN00022; CESP:F38E11.7  
 A; Experimental source: clone F38E11  
 C; Genetics:  
 A; Gene: CESP:F38E11.7  
 A; Intron: 4/50 Position: 4/50

Query Match 40.4%; Score 44; DB 2; Length 284;  
 Best Local Similarity 44.4%; Pred. No. 8.9; Mismatches 8; Conservatve 3; Indels 7; Gaps 0; Matches 8;

Qy 2 PDINPAWYXKGIRPGV 19  
 Db 217 PYIEPMFYALGGLELGR 234

RESULT 7  
T45633 hypothetical protein F13G24.180 - *Arabidopsis thaliana* (mouse/ear cress)  
C;Species: *Arabidopsis thaliana* (mouse/ear cress)  
C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C;Accession: T45623  
R;Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; BE  
submitted to the Protein Sequence Database, December 1999  
A;Reference number: Z23009  
A;Accession: T45623  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 11501 <BEV>  
A;Cross-references: EMBL:AL133421  
A;Experimental source: cultivar Columbia; BAC clone F13G24  
C;Genetics:  
A;Map position: 5  
A;Introns: 64/1; 739/3; 785/2; 1302/2; 1318/3; 1399/2; 1434/2  
A;Note: F13G24.180

Query Match 39.9%; Score 43.5; DB 2; Length 1501;  
Best Local Similarity 34.4%; Pred. No. 57;  
Matches 10; Conservative 5; Mismatches 5; Indels 9; Gaps 2;

Qy 1 TPDPINPAWYXRG----IRPV---GRF 20  
Oy | : | : || | : || | |||  
Db 1184 SPQMAPSWISQYGRKNGLVQPVNNDTGRF 1212

RESULT 8  
C83292 probable glutathione S-transferase PA2821 [imported] - *Pseudomonas aeruginosa* (strain PA  
probable glutathione S-transferase PA2821 [imported] - *Pseudomonas aeruginosa* (strain PA  
C;Species: *Pseudomonas aeruginosa*  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C;Accession: C83292  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
A;Reference number: A82950; MUID:20437337  
A;Accession: C83292  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-309 <NCB>  
A;Cross-references: EMBL:AF025462; PIDN:AAB71002.1; GSPDB:GN00021; CBSP:K10F12.4  
C;Genetics:  
A;Experimental source: strain Bristol N2; clone K10F12  
A;Gene: CEPK10F12.4  
A;Map position: 3  
A;Introns: 31/3; 123/2; 196/3; 239/1

Query Match 39.4%; Score 43; DB 2; Length 232;  
Best Local Similarity 34.4%; Pred. No. 11;  
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 PDINPAWYXRGIRP 16  
Oy | | : | : | : | | |  
Db 130 PDHRAWHLLLGVLPI 144

RESULT 10  
T32376 hypothetical protein K10F12.4 - *Caenorhabditis elegans*  
C;Species: *Caenorhabditis elegans*  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C;Accession: T32376  
R;Wohldmann, J.; Beck, C.  
submitted to the EMBL Data Library, September 1997  
A;Description: The sequence of *C. elegans* cosmid K10F12.  
A;Reference number: Z22157  
A;Accession: T32376  
A;Status: preliminary; translated from GB/EMBL/DBBJ  
A;Molecule type: DNA  
A;Residues: 1-309 <NCB>  
A;Cross-references: EMBL:AF025462; PIDN:AAB71002.1; GSPDB:GN00021; CBSP:K10F12.4  
C;Genetics:  
A;Experimental source: strain Bristol N2; clone K10F12  
A;Gene: CEPK10F12.4  
A;Map position: 3  
A;Introns: 31/3; 123/2; 196/3; 239/1

Query Match 39.4%; Score 43; DB 2; Length 309;  
Best Local Similarity 34.4%; Pred. No. 14;  
Matches 8; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

Qy 2 PDINPAWYXRGIRPGR 19  
Oy | | : | | : | : | | |  
Db 131 PDRSPNWYLPRS--PIGR 146

RESULT 11  
A48197 opsin, ocellar - Atlantic horseshoe crab  
C;Species: *Limulus polyphemus* (Atlantic horseshoe crab)  
C;Date: 17-Jul-1994 #sequence\_revision 17-Jul-1994 #text\_change 13-Aug-1999  
C;Accession: A48197  
R;Smith, W.C.; Price, D.A.; Greenberg, R.M.; Battelle, B.A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 6150-6154, 1993  
A;Title: Opsins from the lateral eyes and ocelli of the horseshoe crab, *Limulus poly*  
A;Reference number: A48197; MUID:93317641  
A;Accession: A48197  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-376 <NCB>  
A;Cross-references: EMBL:103792; PIDN:9156644; PIDN:AA28274.1; PIDN:9156645  
C;Superfamily: vertebrate rhodopsin  
C;Keywords: chromoprotein; G protein-coupled receptor; lipoprotein; photoreceptor; re  
F;318/Binding site: retinal (Lys) (covalent) #status predicted

Query Match 39.4%; Score 43; DB 2; Length 376;

			Best Local Similarity 46.2%; Pred. No. 17; Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
2	PDINPAWXXRG1 14	1 :        :	DIN-----PAWXXRG1 14
40	PRNPPLWISILGV 52		DVNQFFGGPAWMARGL 137
		RESULT 14	
64720	probable amino acid transport protein yaau, sodium-dependent - Escherichia coli	S61046	ARPL protein - yeast ( <i>Saccharomyces cerevisiae</i> )
	Accession: G64720 Species: <i>Escherichia coli</i> R; Pohl, T.M.	N; Alternate names: protein YDL67c	
	Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 26-Aug-1999	C; Species: <i>Saccharomyces cerevisiae</i>	
	Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Roc; Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 29-Oct-1999	C; Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 29-Oct-1999	
	Rose, D.J.; Maurer, B.; Shao, Y.	C; Accession: S61046; S31139; S67719	
	Title: The complete genome sequence of <i>Escherichia coli</i> K-12.	R; Pohl, T.M.	
	Reference number: A64720; MUID:97426617	A; Submitted to the EMBL Data Library, November 1995	
	Accession: G64720 Status: nucleic acid sequence not shown; translation not shown	A; Reference number: S61010	
	Residues: 1-476 <BLAT> A; Accession: S67719	A; Title: Molecular structure and genetic regulation of SFA, a gene responsible for re	
	Residues: 1-476 <BLAT> A; Molecule type: DNA	A; Molecule type: DNA	
	Experimental source: strain K-12; substrain MG655 A; Residues: 1-719 <POW>	A; Reference number: S31138; MUID:93247548	
	Genetics: A; Cross-references: EMBL:Z74215; PIDN:CAA9579.1; PID:91061272	A; Cross-references: EMBL:Z67750; PIDN:91061256; PIDN:CAA9579.1; PID:91061272	
	Gene: SGD:NRPL; ARPL R; Wehner, E.P.; Rao, E.; Brendel, M.	A; Molecule type: DNA	
	Keywords: amino acid transport; transmembrane protein A; Mol. Gen. Genet. 237, 351-358, 1993	A; Cross-references: EMBL:X68020; PIDN:9577609; PIDN:CAA48159.1; PID:9288590	
	10-26/Domain: transmembrane #status predicted <TM1> A; Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992	A; Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992	
	91-107/Domain: transmembrane #status predicted <TM2> R; Pohl, T.M.	A; Pohl, T.M.	
	142-158/Domain: transmembrane #status predicted <TM3> submitted to the EMBL Data Library, November 1995	A; Submitted to the EMBL Data Library, November 1995	
	178-194/Domain: transmembrane #status predicted <TM4> A; Description: The sequence of <i>C. elegans</i> cosmid T13C2.	A; Description: The sequence of <i>C. elegans</i> cosmid T13C2.	
	208-224/Domain: transmembrane #status predicted <TM5> A; Reference number: 218591	A; Reference number: 218591	
	303-319/Domain: transmembrane #status predicted <TM6> A; Accession: T16859	A; Accession: T16859	
	349-365/Domain: transmembrane #status predicted <TM7> A; Status: preliminary; translated from GB/EMBL/DDJB	A; Status: preliminary; translated from GB/EMBL/DDJB	
	391-407/Domain: transmembrane #status predicted <TM8> A; Molecule type: DNA	A; Molecule type: DNA	
	414-430/Domain: transmembrane #status predicted <TM9> A; Residues: 1-156 <D2Z>	A; Residues: 1-156 <D2Z>	
	Query Match 39.4%; Score 43; DB 2; Length 476; A; Cross-references: EMBL:U40030; PIDN:91055164; PIDN:91055165; PIDN:AAA81133.1; CESP:T1	A; Cross-references: MIPS:YDL167c; SGD:S0002326	
	Best Local Similarity 44.4%; Pred. No. 22; Mismatches 8; Conservative 2; Indels 6; Gaps 1;	A; Gene: SGD:YDL167c; SGD:S0002326	
	RESULT 15		
64720	probable amino acid transport protein yaau, sodium-dependent - Escherichia coli	T16859	hypothetical protein T13C2.5 - <i>Caenorhabditis elegans</i>
	Accession: G64720 Species: <i>Escherichia coli</i> C; Species: <i>Caenorhabditis elegans</i>	C; Species: <i>Caenorhabditis elegans</i>	
	Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000 C; Accession: T16859	C; Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000	
	R; Du, Z.	R; Du, Z.	
	Submitted to the EMBL Data Library, November 1995 A; Description: The sequence of <i>C. elegans</i> cosmid T13C2.	Submitted to the EMBL Data Library, November 1995 A; Description: The sequence of <i>C. elegans</i> cosmid T13C2.	
	A; Reference number: 218591 A; Reference number: 218591	A; Reference number: 218591	
	A; Accession: T16859 A; Accession: T16859	A; Accession: T16859	
	A; Status: preliminary; translated from GB/EMBL/DDJB A; Status: preliminary; translated from GB/EMBL/DDJB	A; Status: preliminary; translated from GB/EMBL/DDJB	
	A; Molecule type: DNA A; Molecule type: DNA	A; Molecule type: DNA	
	A; Residues: 1-156 <D2Z> A; Residues: 1-156 <D2Z>	A; Residues: 1-156 <D2Z>	
	Query Match 39.4%; Score 43; DB 2; Length 476; A; Cross-references: EMBL:U40030; PIDN:91055164; PIDN:91055165; PIDN:AAA81133.1; CESP:T1	A; Cross-references: EMBL:U40030; PIDN:91055164; PIDN:91055165; PIDN:AAA81133.1; CESP:T1	
	Best Local Similarity 44.4%; Pred. No. 22; Mismatches 8; Conservative 2; Indels 6; Gaps 1;	A; Gene: CESP:T13C2.5; C; Genetics: C; Genetics: C; Introns: 62/3; 93/3; 124/3; 142/1; 212/1; 293/1; 336/2; 406/3; 437/1; 508/1; 584/1;	

Tue Apr 17 15:46:23 2001

us-09-446-543a-73.rpr

Page 5

Best Local Similarity 45.0%; Pred. No. 60;  
Matches 9; Conservative 2; Mismatches 7;  
Indels 2; Gaps 1;  
Qy 1 TPDINPAWXXRGIRGVGRF 20  
| || | | ; : ||  
Db 975 TTDINSDWYFSR--DINRF 992

Search completed: April 17, 2001, 15:45:59  
Job time: 607 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5 Copyright (C) 1993 - 2000 Compugen Ltd.									
OM protein - protein search, using sw model									
Run on: April 17, 2001, 15:48:53 ; Search time 39.1 Seconds (without alignments) 19.274 Million cell updates/sec									
Title: US-09-446-543A-73 Sequence: 1 TPDINPANWYXKGIRPVGRFXX 22									
Result No.	Score	Query Match	Length	DB ID	Description	Scoring table:	Gapopt 10.0 , Gapext 0.5	searched:	93435 seqs, 34255466 residues
1	105	96.3	87	1	PRRP_HUMAN	BLOSUM62		Total number of hits satisfying chosen parameters:	93435
2	104	95.4	83	1	PRRP RAT	Gapopt 10.0 , Gapext 0.5		Minimum DB seq length: 0	
3	104	95.4	98	1	PRRP_BOVIN	BLOSUM62		Maximum DB seq length: 200000000	
4	46	42.2	676	1	EXL1_HUMAN	Gapopt 10.0 , Gapext 0.5		Post-processing: Minimum Match 0%	
5	43	39.4	376	1	OP52_LIMPO	BLOSUM62		Maximum Match 100%	
6	43	39.4	376	1	CYCR_CHRVL	Gapopt 10.0 , Gapext 0.5		Listing first 45 summaries	
7	43	39.4	383	1	YAAJ_ECOLI	BLOSUM62			
8	43	47.6	476	1	ARP_YEAST	Gapopt 10.0 , Gapext 0.5			
9	43	39.4	719	1	V762_MEJUA	BLOSUM62			
10	41	37.6	342	1	Y5796_MEUTA	Gapopt 10.0 , Gapext 0.5			
11	41	37.6	347	1	CIGR_ARBFU	BLOSUM62			
12	41	37.6	986	1	Y493_MCYTU	Gapopt 10.0 , Gapext 0.5			
13	40	36.7	329	1	CHOD_SRQSQ	BLOSUM62			
14	40	36.7	546	1	streptomyce	Gapopt 10.0 , Gapext 0.5			
15	40	36.7	581	1	POL_MVLRK	BLOSUM62			
16	40	36.7	843	1	POL_MVAK	Gapopt 10.0 , Gapext 0.5			
17	40	36.7	1196	1	POL_MVIRD	BLOSUM62			
18	40	36.7	1196	1	POL_MVIRD	Gapopt 10.0 , Gapext 0.5			
19	39.5	36.2	860	1	VG312_BPR03	BLOSUM62			
20	39	35.9	149	1	ENRN_BPTT	Gapopt 10.0 , Gapext 0.5			
21	39	35.8	360	1	WNF2_HUMAN	BLOSUM62			
22	39	35.8	377	1	OP51_HENSA	Gapopt 10.0 , Gapext 0.5			
23	39	35.8	377	1	OP52_HENSA	BLOSUM62			
24	39	35.8	622	1	PPCC_RAT	Gapopt 10.0 , Gapext 0.5			
25	39	35.8	953	1	SYV_VIBCH	BLOSUM62			
26	39	35.8	962	1	UVRA_METTH	Gapopt 10.0 , Gapext 0.5			
27	38.5	35.3	240	1	P1SC_HELPY	BLOSUM62			
28	38.5	35.3	265	1	UBTE_RCPR	Gapopt 10.0 , Gapext 0.5			
29	38.5	35.3	831	1	DPO1_THREL	BLOSUM62			
30	38.5	35.3	877	1	CAD2_BOVIN	Gapopt 10.0 , Gapext 0.5			
31	38.5	35.3	906	1	CAD2_HUMAN	BLOSUM62			
32	38.5	35.3	906	1	CAD2_MOUSE	Gapopt 10.0 , Gapext 0.5			
33	38.5	35.3	3083	1	POIG_ZYAVR	BLOSUM62			





RL	PROC; NATL ACAD SCI U.S.A 90:6150-6154(1993);	DT	30-MAY-2000 (Rel. 39, Last annotation update)
CC	-1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT	DE	PHOTOSYNTHETIC REACTION CENTER CYTOCHROME C SUBUNIT PRECURSOR.
CC	MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY	PUFC	
CC	LINKED TO CIS-RETINAL.	OS	Chromatium vinosum.
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.	OC	Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;
CC	-1- TISSUE SPECIFICITY: OCULAR CELLS; MEDIAN OCULI.	OC	Allochromatium
CC	-1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY	OX	OX
CC	BE PHOSPHORYLATED (BY SIMILARITY).	RN	NCBI-TAXID=1049;
CC	-1- MISCELLANEOUS: THIS OPSIN HAS AN ABSORPTION MAXIMA AT 530 NM.	RN	[1]
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.	RP	SEQUENCE FROM N.A.
CC	OPSIN SUBFAMILY.	RC	Corson G.P., Nagashima K.V., Matsura K., Sakuragi Y., Ruwanthi W.,
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	RA	Qin H., Allen R., Knaff D.B.;
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	RK	"Primary structure of genes encoding light-harvesting and reaction-
CC	the European Bioinformatics Institute. There are no restrictions on its	RK	center proteins from Chromatium vinosum.";
CC	use by non-profit institutions as long as its content is in no way	RR	Submitted (MAR-1998) to the EMBL/GenBank/DDJB databases.
CC	modified and this statement is not removed. Usage by and for commercial	CC	-1- FUNCTION: THE REACTION CENTER OF PURPLE BACTERIA CONTAIN A
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> )	CC	TIGHTLY BOUND CYTOCHROME MOLECULE WHICH RE-REDUCES THE PHOTO
CC	or send an email to license@sb-sib.ch).	CC	OXIDIZED PRIMARY ELECTRON DONOR.
CC	EMBL: L03792; AAA28274.1; -.	CC	-1- SIMILARITY: HIGH WITH OTHER PHOTOSYNTHETIC REACTION CENTER
CC	PIR: L03187; AAA20499.1; -.	CC	ANCHOR (BY SIMILARITY)
DR	GCRDB; GCR_0585; -.	CC	-1- SIMILARITY: HIGH WITH OTHER PHOTOSYNTHETIC REACTION CENTER
DR	InterPro; IPR00276; -.	CC	CYTOCHROME C SUBUNITS.
DR	InterPro; IPR001391; -.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
DR	InterPro; IPR001760; -.	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
DR	Pfam; PF00001; 7tm_1; 1.	CC	the European Bioinformatics Institute. There are no restrictions on its
DR	PRINTS; PR00237; GPRORHODOPSN.	CC	use by non-profit institutions as long as its content is in no way
DR	PRINTS; PR00238; OPSIN.	CC	modified and this statement is not removed. Usage by and for commercial
DR	PRINTS; PR00578; OPSINLTREYE.	CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> )
DR	PROSITE; PS00337; G-PROTEIN_RBCEP_F1_1; 1.	CC	or send an email to license@sb-sib.ch).
DR	PROSITE; PS00262; G-PROTEIN_RBCEP_F1_2; 1.	CC	PROSITE; PS000345; -.
DR	PROSITE; PS00238; OPSIN_1.	CC	EMBL: AB01181; BAA32742.1.; -.
DR	PROSITE; PS00190; CYTOCHROME_C; 4.	CC	InterPro; IPR000345; -.
KW	Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;	CC	Electron transport; Photosynthesis; Reaction center; Heme;
KW	phosphorylation; G-protein coupled receptor.	CC	Membrane; Lipoprotein; Duplication; Signal;
FT	DOMAIN 1 46 EXTRACELLULAR.	FT	SIGNAL 1 22 BY SIMILARITY.
FT	TRANSMEM 47 71 (POTENTIAL).	FT	CHAIN 23 383 PHOTOSYNTHETIC REACTION CENTER CYTOCHROME
FT	DOMAIN 72 83 CYTOPLASMIC.	FT	C SUBUNIT.
FT	TRANSMEM 84 108 2 (POTENTIAL).	FT	N-ACYL DIGLYCERIDE (BY SIMILARITY).
FT	DOMAIN 109 123 EXTRACELLULAR.	FT	HEME 1 (COVALENT) (BY SIMILARITY).
FT	TRANSMEM 124 143 3 (POTENTIAL).	FT	HEME 1 (COVALENT) (BY SIMILARITY).
FT	DOMAIN 144 162 CYTOPLASMIC.	FT	IRON 1 (HEME AXIAL LIGAND)
FT	TRANSMEM 163 186 4 (POTENTIAL).	FT	(BY SIMILARITY).
FT	DOMAIN 187 210 EXTRACELLULAR.	FT	IRON 1 (HEME AXIAL LIGAND)
FT	TRANSMEM 211 238 5 (POTENTIAL).	FT	(BY SIMILARITY).
FT	DOMAIN 239 274 CYTOPLASMIC.	FT	HEME 2 (COVALENT) (BY SIMILARITY).
FT	TRANSMEM 275 298 6 (POTENTIAL).	FT	HEME 2 (COVALENT) (BY SIMILARITY).
FT	DOMAIN 299 306 EXTRACELLULAR.	FT	IRON 2 (HEME AXIAL LIGAND)
FT	TRANSMEM 307 331 7 (POTENTIAL).	FT	(BY SIMILARITY).
FT	DOMAIN 332 376 CTOPLSMC.	FT	HEME 3 (COVALENT) (BY SIMILARITY).
FT	DISULFID 120 197 BY SIMILARITY.	FT	HEME 3 (COVALENT) (BY SIMILARITY).
FT	BINDING 318 318 (BY SIMILARITY).	FT	IRON 3 (HEME AXIAL LIGAND)
FT	CARBOHYD 17 17 N-LINKED (GLCNAC. ) (POTENTIAL).	FT	(BY SIMILARITY).
FT	CARBONID 193 193 N-LINKED (GLCNAC. ) (POTENTIAL).	FT	HEME 4 (COVALENT) (BY SIMILARITY).
SQ	SEQUENCE 376 AA; 4211 MW; FA9647C40531CBF8 CRC64;	FT	HEME 4 (COVALENT) (BY SIMILARITY).
Query Match 39.4%; Score 43; DB 1; Length 376;	FT	IRON 4 (HEME AXIAL LIGAND)	
Best Local Similarity 46.2%; Pred. No. 8.5; Mismatches 5; Indels 0; Gaps 0;	FT	(BY SIMILARITY).	
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;	FT	SEQUENCE 383 AA; 41522 MW; 96BCD91FF1B9AE7E CRC64;	
RESULT 7 CYCR_CHRV1 STANDARD; PRT; 383 AA.	Query Match 39.4%; Score 43; DB 1; Length 383;	DT	30-MAY-2000 (Rel. 39, Last annotation update)
Qy 2 PDINPAWYXXRG1 14 ID CYCR_CHRV1 STANDARD; PRT; 383 AA.	Best Local Similarity 36.8%; Pred. No. 8.6; Mismatches 9; Indels 0; Gaps 0;	DE	PHOTOSYNTHETIC REACTION CENTER CYTOCHROME C SUBUNIT PRECURSOR.
Db 40 PPMNPLWYSILGV 52 AC 08247; P30143; P30143; P30143; P30143;	Matches 7; Conservative 3; Mismatches 1; Indels 1; Gaps 0;	PUFC	
RESULT 8 YAAJ_ECOLI STANDARD; PRT; 476 AA.	Query Match 1 TPDTINPAWYXXRGTPVGR 19	OS	Chromatium vinosum.
Db 263 TPQTITPAWYAIRHVRDING 281 AC	Best Local Similarity 36.8%; Pred. No. 8.6; Mismatches 9; Indels 0; Gaps 0;	OC	Bacteria; Prokaryotes; Proteobacteria; gamma subdivision; Chromatiaceae;
DT 30-MAY-2000 (Rel. 39, Last sequence update)	Matches 7; Conservative 3; Mismatches 1; Indels 1; Gaps 0;	OC	Allochromatium





CC -!- CATALYTIC ACTIVITY: GTP = 3',5'-CYCLIC GMP + PYROPHOSPHATE.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC  
 DOMAIN OF PROTEIN KINASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X12874; CAIA31367.1; -.  
 DR PIR; S05480; OYURGA;  
 DR InterPro; IPR00719; -.  
 DR InterPro; IPR001828; -.  
 DR Pfam; PF01094; ANF\_Receptor; 1.  
 DR Pfam; PF00069; pkinase; 2.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;  
 CGMP synthesis; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 986 RESACT RECEPTOR.  
 FT DOMAIN 22 507 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 508 528 POTENTIAL.  
 FT DOMAIN 529 986 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 568 836 PROTEIN KINASE LIKE.  
 FT CARBOND 185 185 N-LINKED (GlcNAc. . .) (POTENTIAL).  
 FT CARBOND 361 361 N-LINKED (GlcNAc. . .) (POTENTIAL).  
 FT CARBOHYD 410 410 N-LINKED (GlcNAc. . .) (POTENTIAL).  
 SQ SEQUENCE 986 AA; 111284 MW; B40238A74CCAFCS2 CRC64;

Query Match 37.6%; Score 41; DB 1; Length 986;  
 Best Local Similarity 62.5%; Pred. No. 46;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 PDINPAWY 9  
 Db 475 PDLNPVWH 482

RESULT 13

ID	Y493_MYCTU	STANDARD:	PRT:	329 AA.
AC	Q115C;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	HYPOTHETICAL PROTEIN RV0493C.			
GN	RV0493C OR MTCV20G9-19C.			
OS	Mycobacterium tuberculosis.			
OC	Bacteria; Firmicutes; Actinobacteria; Actinomycetidae; Streptomyces; Actinomycetales; Streptomycetaceae; Streptomyces; Streptomyces sp. (strain SA-COO).			
OC	Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.			
OX	NCBI_TaxID=1773;			
RN	1)			
RE	STRAIN=H37Rv;			
RX	MEDLINE=98195987; PubMed=9634220;			
RP	SEQUENCE FROM N.A.			

Y493\_MYCTU STANDARD: PRT; 329 AA.

Best Local Similarity 42.9%; Pred. No. 23; Length 329; Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 PDINPAWYXXRG 15  
 | ||| : |||  
 Db 90 RAAGPAWFEDIAGVR 103

RESULT 14

ID	CHOD_STRSQ	STANDARD:	PRT:	546 AA.
AC	P12675;			
DT	01-OCT-1989 (Rel. 12, Created)			
DT	01-OCT-1989 (Rel. 12, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	CHOLESTEROL OXIDE PRECURSOR (EC 1.1.3.6) (CHOD).			
GN	CHOA.			
OS	Streptomyces sp. (strain SA-COO).			
OC	Bacteria; Firmicutes; Actinobacteria; Actinomycetidae; Streptomyces; Streptomyces sp. (strain SA-COO).			
OC	Actinomycetales; Streptomycetaceae; Streptomyces.			
OX	NCBI_TaxID=931;			
RN	[1]			
RP	SEQUENCE FROM N.A.			

Y493\_MYCTU STANDARD: PRT; 546 AA.

Best Local Similarity 42.9%; Pred. No. 23; Length 329; Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 PDINPAWYXXRG 15  
 | ||| : |||  
 Db 90 RAAGPAWFEDIAGVR 103

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M31939; AAB26719.1; -.  
 DR PIR; A32260; A32260.  
 DR PDB; 1B4V; 06-JUN-99.  
 DR PDB; 1B8S; 09-FEB-99.  
 DR PDB; 1CBO; 10-MAR-99.  
 DR PDB; 1CC2; 11-MAR-99.

	FT	NON_TER	1	1
	SQ	SEQUENCE	581 AA;	65157 MW;
Query	8D7A38694C8E036E	CRC64;		
Match	7;	Conservative	36.7%;	Score 40;
Matches	7;	Mismatches	41.2%;	DB 1;
Best Local Similarity	36.8%;	Pred. No.	40;	Length 581;
Score	40;	Indels	8;	
Length	581;	Gaps	0;	
DB	556			
Oy	TDDINPAWYXRGIRPV	17		
98	PDKRSSWFKNRTEAPLGSE	116		
Query Match	36.7%;	Score 40;	DB 1;	Length 546;
Best Local Similarity	36.8%;	Pred. No.	38;	
Matches	7;	Mismatches	9;	
Score	40;	Indels	0;	
Length	546;	Gaps	0;	
RESULT	15			
POL_MVRK	STANDARD;	PRT;	581 AA.	
ID				
AC	F31795;			
DT	01-JUL-1993 (Rel. 26; Created)			
DR	15-DEC-1998 (Rel. 37; Last sequence update)			
DE	POLYPROTEIN [CONTAINS: PROTEASE (EC 3.4.23.-); REVERSE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)] (FRAGMENT).			
GN	POL.			
OS	Radiation murine leukemia virus (strain Kaplan).			
OC	Viruses: Retrovirus; Retroviridae; Mammalian type C retroviruses.			
OX	NCBI_TaxID:31689;			
RN	[1]			
RP	SEQUENCE FROM N A.			
RX	MEDLINE:92333703; PubMed:1629969;			
RA	Rassart E.,			
RT	"Determinants of thymotropism in Kaplan radiation leukemia virus and nucleotide sequence of its envelope region.";			
RL	J. Virol. 66:5141-5146(1992)			
CC	-I- PIM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.			
CC	-I- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-POL POLYPROTEIN.			
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A2; ALSO KNOWN AS THE RETROPEPSIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).			
CC	-----			
EMBL: M93052; AAA4625.1; -.				
PIR: A42743; AA2743.				
DR	MEROPS: A02.008; -.			
DR	Interpro; IPR001037; -.			
DR	Interpro; IPR001584; -.			
DR	Interpro; IPR001969; -.			
DR	Interpro; IPR002156; -.			
DR	PFAM: PF0052; Integrase; 1.			
DR	PFAM: PF00075; rnasen; 1.			
DR	PFAM: PF00665; rve; 1.			
DR	PROSITE; PS00141; ASP_PROMEASE; PARTIAL.			
KW	Hydrolase; Transferase; RNA-directed DNA polymerase;			
KW	Aspartyl protease; Endonuclease; Polyprotein.			

FT	NON_TER	1	1
SQ	SEQUENCE	581 AA;	65157 MW;
Query	8D7A38694C8E036E	CRC64;	
Match	7;	Conservative	36.7%;
Matches	7;	Mismatches	41.2%;
Best Local Similarity	36.8%;	Pred. No.	40;
Score	40;	Indels	8;
Length	581;	Gaps	0;
Oy	TDDINPAWYXRGIRPV	17	
Db	556		
Query Match	36.7%;	Score 40;	DB 1;
Best Local Similarity	36.8%;	Pred. No.	40;
Matches	7;	Mismatches	9;
Score	40;	Indels	0;
Length	546;	Gaps	0;
DB	556		
Query Match	36.7%;	Score 40;	DB 1;
Best Local Similarity	36.8%;	Pred. No.	40;
Matches	7;	Mismatches	9;
Score	40;	Indels	0;
Length	546;	Gaps	0;
DB	556		
Search completed: April 17, 2001, 15:48:53			
Job time: 541 sec			

Tue Apr 17 15:46:23 2001

us-09-446-543a-73.rsp

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on:

April 17, 2001, 15:48:13 ; Search time 115.07 Seconds

(without alignments)  
22.409 Million cell updates/sec

Title: US-09-446-543A-73

Perfect score: 109

Sequence: 1 TPDINPAWYXRGIRPVGRFX 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
SPREMBL 15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rabbit:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	82	75.2	117	13 Q9W624
2	49	45.0	790	10 Q9M371
3	46	42.2	54	4 Q9UJF9
4	46	42.2	333	2 Q9PH76
5	45	42.2	465	4 060687
6	46	42.2	540	10 Q9LGZ0
7	45	41.3	767	5 020170
8	44	40.4	284	1 050128
9	44	40.4	419	4 Q9Y276
10	43.5	39.9	1501	10 Q9SD86
11	43	39.4	232	2 Q9RZ21
12	43	39.4	309	5 017234
13	43	39.4	1296	5 022452
14	42.5	39.0	333	2 Q9RJF0
15	42.5	39.0	443	5 Q98T99
16	42	38.5	293	2 Q53B55
17	42	38.5	428	2 Q9PKF7
18	41.5	38.1	345	2 Q9K8V2
19	37.6	2 Q33440		033440 pseudomonas

RESULT

1

Q9W624 PRELIMINARY; PRT; 117 AA.  
ID Q9W624  
AC 09W624;  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)

C-RF AMIDE PRECURSOR.  
Carassius auratus (Goldfish).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OOC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriiformes; Cyprinidae; Cyprininae; Carassius. OX NCBI\_TAXID=7957; RN [1]  
RR SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RA Satake H., Minakata H., Fujimoto M., RT "carassius RFamide (C-RF amide)", RRL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB02024; BA076662.1;  
SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB2038C2B0 CRC64;

## ALIGNMENTS

Query Match	Best Local Similarity	Score	DB ID	Length	PRT
Q9W624 carassius a	60.0%	82	Q9M371	13	117 AA
Q9M371 arabidopsis					
Q9UJF9 homo sapien					
Q9PH76 xylolla fas					
Q96687 homo sapien					
Q9LGZ0 oryza sativa					
Q20170 caenorhabdi					
Q50128 pyrococcus					
Q9Y276 homo sapien					
Q9SD86 arabidopsis					
Q9RZ21 deinococcus					
Q17234 caenorhabdi					
Q22452 caenorhabdi					
Q9RJF0 streptomyce					
Q98T99 caenorhabdi					
Q53B55 spiroplasma					
Q9PKF7 chlamydia m					
Q9K8V2 bacillus ha					
Q33440 pseudomonas					

RESULT	2	PRELIMINARY;	PRT;	790 AA.
Q9M371	ID Q9M371.			
AC 09M371;				
DT 01-OCT-2000 (TREMBLrel. 15, Created)				
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)				
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)				
DE HYPOTHETICAL 87.4 KDA PROTEIN.				
GN F15G16.60.				
OS Arabidopsis thaliana (Mouse-ear cress).				

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TAXID=3702;  
 RN [1] SEQUENCE FROM N.A.  
 RP RA De Haan M., Maarse A.C., Grivell L.A., Mewes H.W., Lemcke K.,  
 RA Mayer K.F.X., Quétier F., Salmonoubat M.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [2] SEQUENCE FROM N.A.  
 RP RA EU Arabidopsis sequencing project;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AL12959; CAB1097\_1; -.  
 KW HYPOTHETICAL PROTEIN  
 SQ SEQUENCE 790 AA; 87375 MW; B22724B75690F30 CRC64;

Query Match 45.0%; Score 49; DB 10; Length 790;  
 Best Local Similarity 47.4%; Pred. No. 7.4; Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
 PRT; 54 AA.

OY 2 PDINPAWXXRGIRGVGRF 20

RESULT 3

ID 09UJF9 PRELIMINARY; PRT; 54 AA.

09UJF9: 01-MAY-2000 (T-EMBL; 13, Created)  
 DT 01-MAY-2000 (T-EMBL; 13, Last sequence update)  
 DE DJ47937\_3 (SUSHI-REPEAT PROTEIN (SRPUL)) (FRAGMENT).  
 GN DJ47937\_3.

OS Homo sapiens (Human);  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TAXID=9606;  
 RN [1] SEQUENCE FROM N.A.

RA Lawlor S.; (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL035608; CAB55682\_1; -.  
 FT NON\_TER 54 54  
 SQ SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;

Query Match 42.2%; Score 46; DB 4; Length 54;  
 Best Local Similarity 43.0%; Pred. No. 1.3; Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;  
 PRT; 54 AA.

OY 1 TPDINPAWXXRGIRP 16

Db 18 TPAVPTWAGSGYVP 33

RESULT 4

09PH76 PRELIMINARY; PRT; 333 AA.

09PH76: 01-OCT-2000 (T-EMBL; 15, Created)  
 DT 01-OCT-2000 (T-EMBL; 15, Last sequence update)  
 DE HYDROXYBENZOATE OCTOPRENYLTRANSFERASE.  
 GN XPO008.  
 OS Xylella fastidiosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 NCBI\_TAXID=2371;  
 RN [1] SEQUENCE FROM N.A.

Query Match 42.2%; Score 46; DB 4; Length 465;

RP STRAIN=945C; RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
 RA Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,  
 RA Barros M.H., Bonacorsi E.D., Bordim-S., Bove J.M., Briones M.R.S.,  
 RA Bueno M.R.P., Camarco A.A., Carrasco D.M., Carrer H.,  
 RA Colaudo N.B., Colombo C., Costa F.F., Costa M.C.R., Costa Neto C.M.,  
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,  
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,  
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.B., Kuramae E.E., Laligret F., Lambais M.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes S.A., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.Y.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Moon D.H., Naegi M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhami A. JR., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peikoto B.R., Pereira G.A.G., Pereira H.A. JR., Pesquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de rosa V.E. JR., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva F.R., Silva W.A. JR.,  
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA Souza A.P., Terenzio M.F., Truffi D., Tsai S.M., Tshuhako M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";  
 RL Nature 406:151-157(2000).  
 DR EMBL; AE003860; AAH82881\_1; -.  
 DR INTERPRO; IPR000537; -.  
 DR PFAM; PF01040; UBA; 1.  
 DR PROSITE; PS00943; UBA; UNKNOWN 1.  
 DR SEQUENCE 333 AA; 37931 MW; ECF3FP4716C962B95 CRC64;

RESULT 5

060687 PRELIMINARY; PRT; 465 AA.

060687: 01-AUG-1998 (T-EMBL; 07, Created)  
 DT 01-AUG-1998 (T-EMBL; 07, Last sequence update)  
 DT 01-MAY-2000 (T-EMBL; 13, Last annotation update)  
 DE SUSHI-REPEAT PROTEIN.  
 GN SRPUL.  
 OS Homo sapiens (Human);  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TAXID=9606;  
 RN [1] SEQUENCE FROM N.A.  
 RP Kurokawa H., Inukai T., Inaba T., Goi K., Chang K.-S., Sinjyo T.,  
 RA Rakeshraw K.M., Naeve C.W., Look T.A.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF060567; ARCI5765\_1; -.  
 DR INTERPRO; IPR000436; -.  
 DR INTERPRO; IPR001128; -.  
 DR PFAM; PF00084; sushi; 3.  
 DR PROSITE; PS0086; sushi; 3.  
 SQ SEQUENCE 465 AA; 52971 MW; 4D752B187FF3EB8 CRC64;

Query Match 42.2%; Score 46; DB 4; Length 465;



		Best Local Similarity	Score	Pred.	No.	Mismatches	Indels	Gaps
OX	NCBI_TAXID-9506;							
RN	[1]	Matches	10;	Conservative	5;	Mismatches	5;	
RP	SEQUENCE FROM N.A.							
RC	TISSUE-BRAIN;							
RX	MEDLINE-9897350; PubMed-9878253;							
RA	petruzzella V., Tiranti V., Fernandez P., Ianna P., Carrozzo R.,							
RA	Zeviani M.;							
RT	"Identification and characterization of human cDNAs specific to BCSL, PET112, SCOL, COX15, and COX11, five genes involved in the formation							
RT	and function of the mitochondrial respiratory chain.;"							
RL	Genomics 54:494 - 504(1998).							
RN	[2]	SEQUENCE FROM N.A.						
RP	TISSUE-BRAIN;							
RC	MEDLINE-95207227; PubMed-8619474;							
RX	A'nderson B., Wentland M.A., Ricafrente J.Y., Liu W., Gibbs R.A.;							
RT	"A 'double adaptor' method for improved shotgun library							
RT	construction.";							
RL	Anal. Biochem. 236:107-113(1996).							
RN	[3]	SEQUENCE FROM N.A.						
RP	SEQUENCE FROM N.A.							
RC	TISSUE-BRAIN;							
RX	MEDLINE-97264341; PubMed-9110174;							
RA	Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;							
RT	"Large-scale concatenation cDNA sequencing.;"							
RL	Genome Res. 7:353-358(1997).							
DR	EMBL; AF020849; AAD06638.1; -;							
DR	AF038195; AAB97365.1; -;							
DR	INTERPRO; IPR001939; -;							
DR	PFAM; PF00004; AAA; 1.							
KW	Hypothetical protein							
SQ	SEQUENCE 419 AA; 47534 MW; 7F0F98BA62F2CBBB CRC64;							
Query Match	40.4%; Score 44; DB 4; Length 419;							
Best Local Similarity	70.0%; Pred. No. 25;							
Matches	7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;							
Qy	5 NPAWYXXARGI 14							
Db	211 NPKWYDORG1 220							
RESULT	10							
O9SDB6	PRELIMINARY; PRT; 1501 AA.							
AC	O9SDB6;							
DT	01-MAY-2000 (TREMBLrel. 13, Created)							
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)							
DE	HYPOTHETICAL 164.4 KDA PROTEIN.							
GN	F13G24.180.							
OS	Arabidopsis thaliana (Mouse-ear cress).							
OC	Eukaryota; Viriciliophytida; Embryophyta; Tracheophyta; Spermatophyta;							
OC	Magnoliophyta; eudicots; Rosidae; eurosids II;							
OC	Brassicales; Brassicaceae; Arabidopsis.							
NCBI-TAXID=3702;								
RN	[1]	SEQUENCE FROM N.A.						
RP	SEQUENCE FROM N.A.							
RA	Bevan M., Van Der Schueren J., Chuang Y.J., Voet M., Robben J., Volkert G., Bancroft I., Mewes H.W., Lemcke K., Mayer K.F.X.;							
RL	Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.							
RN	[2]	SEQUENCE FROM N.A.						
RP	EU Arabidopsis sequencing project;							
RL	Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.							
RN	[3]	SEQUENCE FROM N.A.						
RP	EU Arabidopsis sequencing project;							
RL	Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.							
RN	[4]	SEQUENCE FROM N.A.						
RP	SEQUENCE FROM N.A.							
RESULT	11							
O9RZ21	PRELIMINARY; PRT; 232 AA.							
ID	O9RZ21;							
AC	O9RZ21;							
DT	01-MAY-2000 (TREMBLrel. 13, Created)							
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)							
DR	HYPOTHETICAL 25.0 KDA PROTEIN.							
GN	DRA0122.							
OS	Deinococcus radiodurans.							
OC	Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus							
OX	NCBI_TAXID-1299;							
RN	[1]	SEQUENCE FROM N.A.						
RP	SEQUENCE FROM N.A.							
RC	STRAIN-R1;							
RX	MEDLINE-20036895; PubMed-10567266;							
RA	White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D., Dodson R.J., Haft D.H., Gdin M.L., Nelson W.C., Richardson D.L., Moffat K.S., Qin H., Jiang L., Pamphilis W., Crosby M., Shen M., Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C., Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R., Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C., Fraser C.M.;							
RA	"Genome Sequence of the Radioresistant Bacterium Deinococcus radiodurans R1"; Radiotrans R1; Science 285:1571-1577(1999). DR EMBL; AE001862; AAC12317.1; -.							
DR	TIGR; DRA0132; -.							
KW	Hypothetical protein.							
SQ	SEQUENCE 232 AA; 24979 MW; A044FA2F38435Dn7 CRC64;							
Query Match	39.4%; Score 43; DB 2; Length 232;							
Best Local Similarity	46.4%; Pred. No. 20;							
Matches	7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;							
Qy	2 PDINPAWYXXRGIP 16							
Db	130 PDHRAWHLKGVL 144							
RESULT	12							
O17234	PRELIMINARY; PRT; 309 AA.							
ID	O17234;							
AC	O17234;							
DT	01-JAN-1998 (TREMBLrel. 05, Created)							
DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)							
DE	01-MAY-2000 (TREMBLrel. 13, Last annotation update)							
DE	K10F12.4 PROTEIN.							
GN	K10F12.4.							
OS	Caenorhabditis elegans.							
OC	Bukarota; Metazoa; Nematoidea; Chromadorea; Rhabditida; Rhabditoidae; Peidorinae; Caenorhabditidae.							
OC	Rhabditidae; Peidorinae; Caenorhabditidae.							
OX	NCBI_TAXID-629;							
RN	[1]	SEQUENCE FROM N.A.						
RP	SEQUENCE FROM N.A.							
RC	STRAIN-BRISTOL N2;							
RX	MEDLINE-94150718; PubMed-7906398;							
RA	Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirstein J., Lalister N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rikken L., Rocra A., Saunders D., Shownkeen R.,							

RA	Smalldon N., Smith A., Sonhammer E., Staden R., Sulston J., Waterston R.	DR	INTERPRO; IPR002049; -.
RA	Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.; "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegens."; Nature 368:32-38(1994).	DR	INTERPRO; IPR002350; -.
RT		DR	PFAM; PF0050; Kazal; 9.
RT		DR	PFAM; PF0053; laminin_EGF; 2.
DR		DR	PFAM; PF0054; laminin_G; 1.
DR		DR	PRINTS; PR0290; KAZALINHBT.
DR		DR	PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR		DR	PROSITE; PS01248; LAMININ_TYPE_EGF; 1.
DR		DR	SEQUENCE; 1296 AA; 145178 MW; 05094BC185839690 CRC64;
RN	[3] Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.	RN	SEQUENCE FROM N.A.
RP		RC	STRAIN=BRISTOL N2;
RA		RC	Waterston R.;
RL		RC	Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR		DR	EMBL; AF025462; AABP1002.1; -.
DR		DR	INTERPRO; IPR00521; -.
DR		DR	PFAM; PF00053; laminin_EGF; 2.
SQ	SEQUENCE 309 AA; 35021 MW; FA2091931D8AD2DC CRC64;	SQ	SEQUENCE FROM N.A.
RESULT	13	Query Match	39.4%; Score 43; DB 5; Length 309;
ID	022452	Match	Best Local Similarity 44.4%; Pred. No. 27; Mismatches 8; Conservative 3; MisMatches 5; Indels 2; Gaps 1;
AC	022452:	PRT	PRELIMINARY; PRT; 1296 AA.
OY	2	DT	PDINPAWXXRGIRPVGR 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
Db	131 PDRSPNWLRS--PIGR 146	DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)
OS		DE	SIMILAR TO AGRIN AND FOLLISTATIN. T13C2.5.
OC		OS	Baculovirus; Metazoa; Chromadorea; Rhabditida; Rhabditoidea; Caenorhabditis elegans.
OC		OC	Rhabditida; Peletierinae; Caenorhabditis.
OX		OC	NCBI_TaxID=6239;
RN	[1] SEQUENCE FROM N.A.	OX	Actinomycetales; Streptomyceae; Streptomycetaceae; Streptomyces.
RX	MEDLINE=94150718; PubMed=7906398;	RN	NCBI_TaxID=1902;
RA	Wilson R., Alnæsough R., Anderson K., Baynes C., Berks M., Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kister N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Riffen L., Roopra A., Saunders D., Shownkeen R., Smalton N., Smith A., Sonhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;	RA	SEQUENCE FROM N.A.
RA	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans."; Nature 368:32-38(1994).	RA	STRAIN=A3(2); Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RP		RA	Seeger K.J., Harris D.; [2]
RC		RA	SEQUENCE FROM N.A.
RN	[2] SEQUENCE FROM N.A.	RA	STRAIN=A3(2);
RP		RA	Parthill J., Barrell B.G., Rajandream M.A.; [3]
RC		RA	Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN		RA	SEQUENCE FROM N.A.
RP		RA	STRAIN=A3(2); Reddenbach M., Kieser H.M., Denapaitte D., Eichner A., Cullum J., Kinashi H., Hopwood D.A.;
RC		RA	"A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome."; Mol. Microbiol. 21:77-96(1996); EMBL; AL121746; CAB57411.1; -.
RN		KW	Hypothetical protein.
RP		SQ	SEQUENCE 333 AA; 36312 MW; 884FF2861837FD76 CRC64;
RA		Query Match	39.0%; Score 42.5; DB 2; Length 333;
RT		Match	Best Local Similarity 32.3%; Pred. No. 35; Mismatches 10; Conservative 2; MisMatches 8; Indels 11; Gaps 1;
RL		PRT	PRELIMINARY; PRT; 443 AA.
RN	[3] Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.	QY	1 PDINPAW-----YXXRGIRPVGRF 20
RP		Db	19 TPWEEPAWPTEAIGWIEARLAAGHSRPTGRW 49
RA	SEQUENCE FROM N.A.	RESULT	15
RL	Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.	ID	Q19879
DR		AC	Q19879;
DR		DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DR		DT	01-NOV-1996 (TREMBLrel. 01, Last annotation update)
DR		DT	01-JAN-1999 (TREMBLrel. 09, Last annotation update)

DE F28D1.8 PROTEIN.  
 GN F28D1.8.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderrinae; Caenorhabditis.  
 OX NCBI-TAXID=6239;  
 RN [1]  
 RP Baynes C.;  
 RL Submitted (APR-1996) to the EMBL/genBank/DDJB databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906598;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Cossey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Montimore B., O'callaghan M.,  
 RA Parsons J., Percy C., Riffen L., Roopra A., Saunders D., Showkeen R.,  
 RA Smalton N., Smith A., Sounhamer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Spratt J., Wohldman P.,  
 RT 2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-39(1994).  
 DR EMBL: 270084; CAA94603.1; -;  
 SQ SEQUENCE 443 AA; 48865 MW; D9137BAACCEF575E CRC64;

Query Match 39.0%; Score 42.5; DB 5; Length 443;  
 Best Local Similarity 37.5%; Pred. No. 47;  
 Matches 9; Conservative 3; Mismatches 3; Indels 9; Gaps 1;

Qy	4	INPAWYXXR-----GIRPVG 18
Db	53	INPSWFDFWRVSVSIDGHGLGIWPIG 76

Search completed: April 17, 2001, 15:48:13  
 Job time: 566 sec